Biomechanical Parameter Fingerprint in the Mucosal Wave
Power Spectral Density

P. Gómez, J. I. Godino, F. Díaz, A. Álvarez, R. Martínez, V. Rodellar

Facultad de Informática, Universidad Politécnica de Madrid
Campus de Montegancedo, s/n, 28660 Boadilla del Monte, Madrid, Spain
pedro@pino.datsi.fi.upm.es

Abstract
The importance of mucosal wave detection and estimation has been stressed in literature regarding the automatic classification and recognition of larynx pathologies from voice records. Using a new estimation method of the mucosal wave correlate and simulation results from a 2-mass model of the vocal folds the present paper shows that the main fingerprints found in the power spectral density of the mucosal wave correlate are directly related to the biomechanics of the system. These findings open the door to the non-invasive estimation of biomechanical parameters of the vocal folds directly from voice records, thus easing the task of automatic pathologic classification and recognition.

1. Introduction
In what follows the term mucosal wave will refer to the travelling wave effect taking place in the vocal cords due to the distribution of masses on the body cover and related tissues, and the term mucosal wave correlate (MWC) will be used for the influence of mucosal wave on the overall pattern of the glottal aperture, appearing as a superimposed ringing on its reconstructed trace. The presence and specific pattern of the mucosal wave is of most importance in establishing the presence of certain pathologies in the vocal folds [8]. The MWC may be seen as a higher order vibration regime of the vocal folds, once the average main movement or first regime has been removed, as will be shown later. It is well know to larynx pathology experts that the power spectral density of voice hides precise hints to the presence of certain voice pathologies, as for example nodules [7]. The distribution of the power spectral density of voice on frequencies and the relation among the peaks in the envelope, and lower harmonics of the spectral density have been used with this purpose, these being some of the characteristics used (amplitudes given in dB):
- Difference in amplitude between the first harmonic and the harmonic closest to the first formant.
- Difference in amplitude between the first harmonic and the harmonic closest to the third formant.
- Difference in amplitude between the harmonic closest to the first formant and the harmonic closest to the third formant.
- Difference in amplitude between the first two harmonics.

This reveals that there must be a connection among the power spectral distribution and the presence of pathology. To study this the present research proceeds in two associated directions to deepen in the idea of using the voice power spectral density in the detection of larynx pathologies. It is first shown that the mucosal wave correlate is a better instrument for pathology detection than the raw spectrum of voice. This is so because the influence of the vocal tract and of the first order of vibration of the vocal cords have been removed from the MWC, which will be only influenced by vocal cord dynamics. The second direction of exploration is devoted to study the transfer function between mechanical forces acting on the vocal folds due to air pressure and the average speed of the cords during phonation (vocal cord mechanical trans-admittance), as a good representative of vocal cord dynamics. It is shown that this transfer function is directly responsible for the main features to be found in the spectral density of the mucosal wave. Once this relation is stated the intention of the research is to find methods to directly estimate the biomechanical parameters involved in the cord mechanical trans-admittance which may lead to determine the distribution of masses and inter-elasticities in a 2-mass model approaching the behavior of the real vocal fold system, thus opening the possibility of pathology detection and classification by automatic means.

2. Cord Mechanical Trans-Admittance
Under the purpose of the present work a model including the most relevant features of the vocal cord system have been used and adapted [4]: 2-mass asymmetrical modelling, non-linear coupling between mass movement and glottal aperture, cord collision effects and non-linearities taken into account, deficient closure effects, lung flux excitation and vocal tract coupling effects. This is a modification of a well-known model through literature on the topic [1], [2]. The main characteristics of the model may be seen in Figure 1.

Figure 1. Schematic structure of the two-mass model.

To start the study a version of the vocal cord 2-mass model as given in [6] has been implemented in MATLAB® [4], its main features being: 2-mass asymmetric modelling, non-linear coupling between mass movement and glottal aperture, cord collision effects, non-linearities and defective closure effects.
taken into account, lung flux excitation and vocal tract coupling. The parameters of the model are the lumped masses (2 per cord) \(M_{1l}\) and \(M_{2l}\) (left cord), \(M_{1r}\) and \(M_{2r}\) (right cord), the elastic parameters \(K_{1l}\) and \(K_{2l}\) (relative to reference) and \(K_{12l}\) (inter-coupling), and their respective ones for the right cord: \(K_{1r}\), \(K_{2r}\) and \(K_{12r}\). The dynamic equations of the model are a set of four integro-differential equations, the two for the right-hand cord being:

\[
\begin{align*}
fs1 &= v_{r1}^{2}R_{r1} + M_{1r} \frac{dv_{r1}}{dt} + K_{1r} \int v_{r1}^{2} dt + K_{12r} \int (v_{r1} - v_{r2}) dt \quad (1) \\
fs2 &= v_{r2}^{2}R_{r2} + M_{2r} \frac{dv_{r2}}{dt} + K_{2r} \int v_{r2}^{2} dt + K_{12r} \int (v_{r2} - v_{r1}) dt \quad (2)
\end{align*}
\]

It is of most importance for our study to consider the behavior of the model in the frequency domain, especially in associating the forces \(f_{s1}\) and \(f_{s2}\) acting on both masses of the same cord (as for instance \(M_{1r}\) and \(M_{2r}\)) with the observable mass velocities \(v_{r1}\) and \(v_{r2}\) in the electromechanical equivalent circuit given in Figure 2, where mechanical forces play the role of electromotive forces and velocities play the role of currents.

![Figure 2](https://via.placeholder.com/150)

**Figure 2.** Equivalent electromechanical circuit of the right cord in the two-mass model.

This system may be solved using classical System Theory, as Laplace methods for transient regimes, or Heaviside equivalents for the sinusoidal permanent case. Assuming null initial conditions at rest for both cords, the dynamic equations could be described in the domain of Laplace as

\[
\begin{align*}
F_{s1}(s) &= Z_{11r}(s) V_{r1}(s) + Z_{12r}(s) V_{r2} \quad (3) \\
F_{s2}(s) &= Z_{21r}(s) V_{r1}(s) + Z_{22r}(s) V_{r2} \quad (4)
\end{align*}
\]

where the parameters of the system will be given as

\[
\begin{align*}
Z_{11r}(s) &= sM_{1r} + R_{1r} + s^{-1}(K_{1r} + K_{12r}) \quad (5) \\
Z_{12r}(s) &= -s^{-1}K_{12r} \quad (6) \\
Z_{21r}(s) &= sM_{1r} + R_{1r} + s^{-1}(K_{1r} + K_{12r}) \quad (7)
\end{align*}
\]

Our interest will be to know the response of the two-cord system to an impulse excitation in one of the masses when the other is at rest (this being an ideal situation which does not meet a practical case), with the purpose of gaining more insight about the system behavior. For such we will impose the condition \(F_{s2}=0\) under a permanent sinusoidal regime \((s=j\omega)\). Under such circumstances a relation between the active force \(F_{s1}(\omega)\) on the supraglottal lips and the associated velocity \(V_{r1}(\omega)\) on the supraglottal mass \(M_{1r}\) will take the form of a **trans-admittance transfer function** its square modulus being

\[
\left| V_{r1}(\omega) \right|^2 = \frac{a_{1r}\omega^6 + a_{2r}\omega^4 + a_{1r}\omega^2}{b_{1r}\omega^8 + b_{2r}\omega^6 + b_{3r}\omega^4 + b_{4r}\omega^2 + b_{5r}} \quad (8)
\]

where \(\{a_{ij}\}\) and \(\{b_{ij}\}\) are explicit functions of the biomechanical parameters of the model \(M_{1r}, M_{2r}, R_{1r}, R_{2r}, K_{1r}, K_{2r}, K_{12r}\), not shown here for the sake of brevity. To better understand the behavior of the **vocal cord mechanical trans-admittance** a plot of its modulus and phase in the frequency domain from 0-10,000 Hz is given in Figure 3.

![Figure 3](https://via.placeholder.com/150)

**Figure 3.** Modulus and phase of the trans-admittance for specific values of the biomechanical parameters \((M_{1r}, M_{2r}, R_{1r}, R_{2r}, K_{1r}, K_{2r}, K_{12r})\).

The function in (8) presents four maxima for real frequencies, two in the positive part and two in the negative part of the \(\omega\) axis, due to the minima in the denominator. On its turn the numerator will present a zero for \(\omega=0\), and one minimum on each side of the \(\omega\) axis.

### 3. Results

It is now time to study if the behavior of the **vocal cord mechanical trans-impedance** can be observed on the power spectral density of any of the correlate signals involved in glottal dynamics, for which we present in Figure 4 the input voice and the reconstructed mucosal wave correlate of a male subject uttering the vowel /a/.

![Figure 4](https://via.placeholder.com/150)

**Figure 4.** Top: Derivative of the glottal pulse (glottal aperture) obtained after inverse filtering and leveling. Middle: Average excursion of the glottal aperture or SLRC. Bottom: Mucosal Wave Correlate or FSRC.

In the upper template the glottal aperture or derivative of the glottal pulse \(dU_g/dt\) is presented, obtained by inverse filtering.
and leveled to a common ground line as described in [3]. For the purposes of obtaining the mucosal wave correlate a smoothing technique is used [5] which renders the average movement of the glottal aperture as shown in the middle template, reflecting the average slow moving amplitude of the cord lips, this signal being referred to as the Slow and Large Range Component (SLRC). The mucosal wave correlate is obtained from the difference between these two traces as represented in the bottom template, exposing the rapid and small relative movement between the twocord masses in the model, being referred to as the Fast and Small Range Component (FSRC). The associated power spectral densities of these three traces are given in Figure 5.

Figure 5. Power spectral densities associated to the traces in Figure 4.

It may be seen that the spectral behavior of the derivative of the glottal pulse (glottal aperture) in the upper template shows a decay with frequency as should be expected, with notches which could be associated with the singularities of the vocal cord trans-admittance, but this behavior becomes more evident when the influence of the average component (middle) or SLRC is removed, the result being plotted in the bottom template. The main characteristics of the vocal cord trans-admittance may be appreciated on this signal, as two hunchbacked maxima separated by a sharp notch. As this behavior may not resemble very much that shown in Figure 6, which keeps much more similarities with the one in the bottom of Figure 5.

Figure 6. Square modulus of the vocal cord trans-admittance for a set of values of the biomechanical parameters matching the behavior of the mucosal wave correlate power spectral density in Figure 5 (bottom).

Especially relevant for this adjustment was the ratio \( r_{12} = M_2/M_1 \), which is responsible of the relative positions of the two maxima and the interspersed minimum. Other biomechanical parameters had their influence in the respective positions of the maxima, as the ratio between the elasticities \( r_{12} = K_2/K_1 \).

4. Discussion

At this point it will be of most interest to evaluate the presence of singularities in the mucosal wave correlate obtained by the procedure described above. For such the traces produced by 16 normal speakers (supposedly free from larynx pathology) were processed, half of them males and half females, uttering the vowel /a/ for about 2 sec. The utterances were clipped for 0.2 sec. around their middle part (to avoid onset and decay effects), and the singularity parameters were evaluated over an average of the associated number of cycles present in the trace, in the range 18-40, depending on each case. The results are shown in Table 1.

<table>
<thead>
<tr>
<th>Subj.</th>
<th>Sex</th>
<th>Val. 1st max. (dB)</th>
<th>freq. (Hz)</th>
<th>Val. min. (dB)</th>
<th>freq. (Hz)</th>
<th>Val. 2nd max. (dB)</th>
<th>freq. (Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>#6.4</td>
<td>M</td>
<td>3.62</td>
<td>439</td>
<td>-36.95</td>
<td>769</td>
<td>-6.97</td>
<td>989</td>
</tr>
<tr>
<td>#6.5</td>
<td>M</td>
<td>12.38</td>
<td>467</td>
<td>-10.65</td>
<td>747</td>
<td>1.76</td>
<td>841</td>
</tr>
<tr>
<td>#6.7</td>
<td>M</td>
<td>8.04</td>
<td>344</td>
<td>-3.01</td>
<td>574</td>
<td>6.41</td>
<td>804</td>
</tr>
<tr>
<td>#6.8</td>
<td>M</td>
<td>-2.14</td>
<td>370</td>
<td>-12.99</td>
<td>493</td>
<td>-7.44</td>
<td>617</td>
</tr>
<tr>
<td>#6.8</td>
<td>M</td>
<td>10.18</td>
<td>336</td>
<td>-15.58</td>
<td>504</td>
<td>8.36</td>
<td>588</td>
</tr>
<tr>
<td>#7.1</td>
<td>M</td>
<td>2.41</td>
<td>357</td>
<td>-27.17</td>
<td>833</td>
<td>-11.16</td>
<td>1071</td>
</tr>
<tr>
<td>#6.E</td>
<td>M</td>
<td>11.85</td>
<td>434</td>
<td>-12.28</td>
<td>869</td>
<td>-2.15</td>
<td>978</td>
</tr>
<tr>
<td>#6.F</td>
<td>M</td>
<td>5.43</td>
<td>404</td>
<td>-20.48</td>
<td>707</td>
<td>-12.33</td>
<td>909</td>
</tr>
<tr>
<td>#0.4</td>
<td>F</td>
<td>-1.76</td>
<td>416</td>
<td>-18.92</td>
<td>833</td>
<td>-15.09</td>
<td>972</td>
</tr>
<tr>
<td>#0.7</td>
<td>F</td>
<td>-1.72</td>
<td>638</td>
<td>-24.72</td>
<td>1063</td>
<td>-21.32</td>
<td>1276</td>
</tr>
<tr>
<td>#1.2</td>
<td>F</td>
<td>3.42</td>
<td>769</td>
<td>-20.05</td>
<td>1346</td>
<td>-14.87</td>
<td>1730</td>
</tr>
<tr>
<td>#1.3</td>
<td>F</td>
<td>-7.10</td>
<td>612</td>
<td>-30.62</td>
<td>1020</td>
<td>-9.65</td>
<td>1632</td>
</tr>
<tr>
<td>#2.6</td>
<td>F</td>
<td>4.50</td>
<td>689</td>
<td>-19.11</td>
<td>2096</td>
<td>-12.16</td>
<td>2413</td>
</tr>
<tr>
<td>#2.9</td>
<td>F</td>
<td>2.08</td>
<td>888</td>
<td>-23.90</td>
<td>2222</td>
<td>-18.67</td>
<td>2444</td>
</tr>
<tr>
<td>#2.B</td>
<td>F</td>
<td>-6.47</td>
<td>784</td>
<td>-24.52</td>
<td>1176</td>
<td>-17.75</td>
<td>1568</td>
</tr>
</tbody>
</table>

Table 1. Results after detecting the singularities found in the mucosal wave correlate for 16 normal subjects.

The first conclusion from the data in the table, is that the variability of data is relatively important even for normal voices. Female voices show positions for the singularities at higher frequencies than male voices, as should be expected. Besides, female voices show a less sharper notch than male voices. To help in comparing the absolute values given in Table 1, a normalized set of values is given in Table 2. From left to right the third column gives the depth of the notch relative to the positions of the maxima, as the ratio between the elasticities \( r_{12} = K_2/K_1 \). The fourth one gives the position of the notch relative to the position of the first maximum \( f_n \). The fifth one gives the height of the second maximum relative to the first one. The sixth one gives the relative position of the second minimum relative to the first one \( f_{21} \). The seventh column gives the notch slenderness factor, which is defined as

\[
N_{sf} = \frac{|Y_n|^2}{f_{21} - f_{n}} \tag{9}
\]

The eighth column (right-most) gives the normalized position of the notch relative to the positions of the two maxima, defined as

\[
N_{sf} = \frac{|Y_n|^2}{f_{21} - f_{n}} \tag{9}
\]
$p_{im} = 2 \frac{f_{im}}{f_{r1} + 1}$ \hspace{1cm} (10)

It may be seen in this case that although the slenderness of the notch is rather variable from one subject to others (being larger in general for males than for females), the relative position of the notch to the central position between the maxima is more regular for most of the subjects.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>#6.4</td>
<td>M</td>
<td>-40.57</td>
<td>1.75</td>
<td>-10.59</td>
<td>2.25</td>
<td>32.38</td>
<td>1.08</td>
</tr>
<tr>
<td>#6.5</td>
<td>M</td>
<td>-23.03</td>
<td>1.60</td>
<td>-10.62</td>
<td>1.80</td>
<td>28.76</td>
<td>1.14</td>
</tr>
<tr>
<td>#6.7</td>
<td>M</td>
<td>-11.05</td>
<td>1.67</td>
<td>-1.63</td>
<td>2.34</td>
<td>8.26</td>
<td>1.00</td>
</tr>
<tr>
<td>#6.8</td>
<td>M</td>
<td>-10.85</td>
<td>1.33</td>
<td>-5.30</td>
<td>1.67</td>
<td>16.25</td>
<td>1.00</td>
</tr>
<tr>
<td>#6.B</td>
<td>M</td>
<td>-31.00</td>
<td>1.50</td>
<td>-7.06</td>
<td>1.75</td>
<td>41.33</td>
<td>1.09</td>
</tr>
<tr>
<td>#7.1</td>
<td>M</td>
<td>-29.58</td>
<td>2.33</td>
<td>-13.57</td>
<td>3.00</td>
<td>14.79</td>
<td>1.17</td>
</tr>
<tr>
<td>#6.E</td>
<td>M</td>
<td>-24.13</td>
<td>2.00</td>
<td>-14.00</td>
<td>2.25</td>
<td>19.25</td>
<td>1.23</td>
</tr>
<tr>
<td>#6.F</td>
<td>M</td>
<td>-25.91</td>
<td>1.75</td>
<td>-17.76</td>
<td>2.25</td>
<td>20.73</td>
<td>1.08</td>
</tr>
<tr>
<td>#0.F</td>
<td>F</td>
<td>-17.16</td>
<td>2.00</td>
<td>-13.33</td>
<td>2.34</td>
<td>12.84</td>
<td>1.20</td>
</tr>
<tr>
<td>#0.7</td>
<td>F</td>
<td>-23.00</td>
<td>1.67</td>
<td>-19.60</td>
<td>2.00</td>
<td>23.00</td>
<td>1.11</td>
</tr>
<tr>
<td>#1.2</td>
<td>F</td>
<td>-23.47</td>
<td>1.75</td>
<td>-18.29</td>
<td>2.25</td>
<td>18.78</td>
<td>1.08</td>
</tr>
<tr>
<td>#1.3</td>
<td>F</td>
<td>-23.52</td>
<td>1.67</td>
<td>-2.55</td>
<td>2.67</td>
<td>14.11</td>
<td>0.91</td>
</tr>
<tr>
<td>#2.6</td>
<td>F</td>
<td>-23.61</td>
<td>3.00</td>
<td>-16.66</td>
<td>3.50</td>
<td>9.44</td>
<td>1.33</td>
</tr>
<tr>
<td>#2.9</td>
<td>F</td>
<td>-25.98</td>
<td>2.50</td>
<td>-20.75</td>
<td>2.75</td>
<td>14.83</td>
<td>1.33</td>
</tr>
<tr>
<td>#2.A</td>
<td>F</td>
<td>-17.30</td>
<td>1.67</td>
<td>-7.30</td>
<td>2.00</td>
<td>17.27</td>
<td>1.11</td>
</tr>
<tr>
<td>#2.B</td>
<td>F</td>
<td>-18.05</td>
<td>1.50</td>
<td>-11.28</td>
<td>2.00</td>
<td>18.05</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Table 2. Normalized results for peak slenderness and position.

The variability found is due to the specific phonation characteristics and style of each speaker, and does not put an inconvenience to the study, as the main challenge is how to use the positions and values of the singularities to determine the values of the biomechanical parameters of the associated 2-mass model. A possible strategy for biomechanical parameter estimation should be to match the 6 estimations of the singular values automatically against the biomechanical parameters of the 2-mass model adjusting these using a steering engine. In this sense several propositions may be put into work, one of them based in fitting the singular values estimated from the vocal cord trans-admittance moving the biomechanical models to approximate the estimations from the mucosal wave correlate by means of the Jacobian Function of the vocal cord trans-admittance, which may be pre-calculated and stored on a table. Other techniques could be based on pseudo-inverse matching, or even in neural networks. As the number of observable parameters (3 values and 3 positions) is lower than the number of biomechanical parameters to be adjusted (2 masses, 2 mechanical resistances and 3 elasticities), some extra observables could be measured, as for example the bandwidth of the first two singularities (maximum and minimum) strongly related with the parameters $R_1$ and $R_2$, although this could increased the sensitivity of the system to measurement errors, thus reducing reliability. One added difficulty will be the detection of unbalances between the left and right cords, which would be especially apparent in certain kinds of pathologies. This will be handled by processing differentially every two neighbor cycles.

5. Conclusions

Through the present paper it has been shown that specific characteristics present in the power spectral density of the mucosal wave correlate obtained from the derivative of the glottal pulse can be due to the singularities present in the transfer function of the vocal cord trans-admittance relating the force exerted on one of the cord masses and the velocity of the same mass in the frequency domain. These singularities appear as two maxima or peaks in the transfer function and an interspersed minimum between them. As there is a clear connection between the singularity positions and size and the biomechanical parameters associated to the 2-mass model supporting the general form of the vocal cord trans-admittance, it may be expected that the biomechanical parameters of the model could be estimated from the direct measurement of the positions, values and bandwidths associated to the singularities, thus opening the door to non-invasive characterization of the vocal cord system with applications in pathology characterization.

6. Acknowledgments

This research is carried out under the following projects of Programa Nacional de las Tecnologías de la Información y las Comunicaciones from the Ministry of Science and Technology of Spain: TIC-2002-0273, TIC2003-08956-C02-00, TIC2003-08756.

7. References